fw

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/685,585	10/16/2003	Daben Liu	02-4018	5147

25537 7590 10/05/2007

VERIZON
PATENT MANAGEMENT GROUP
1515 N. COURTHOUSE ROAD
SUITE 500
ARLINGTON, VA 22201-2909

EXAMINER	
SIEDLER, DOROTHY S	
ART UNIT	**PAPER NUMBER**
2626	

NOTIFICATION DATE	DELIVERY MODE
10/05/2007	ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

patents@verizon.com

Office Action Summary	Application No.	Applicant(s)	
	10/685,585	LIU ET AL.	
	Examiner	Art Unit	
	Dorothy Sarah Siedler	2626	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on 06 July 2007.

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) 1-12 and 14-31 is/are pending in the application.

4a) Of the above claim(s) 13 is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) 1-12 and 14-31 is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on 10-16-03 is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)

2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)

3) ☐ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _____.

4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .

5) ☐ Notice of Informal Patent Application

6) ☐ Other: _____.

DETAILED ACTION

This action is in response to the amendment filled July 16, 2007. Claims 1-31 are pending; claim 13 is canceled and claims 1,2,10,16,19 and 28 are amended.

Response to Arguments

The applicant's arguments with regard the 35 U.S.C 101 rejection of claims 1-2 and 28-31 are persuasive. Therefore the rejection is withdrawn.

Applicant's arguments filed July 16, 2007 have been fully considered but they are not persuasive.

Applicant asserts that, "As Liu is generally directed to detecting the number of parameters analyzed to determine speaker changes quickly, Liu not only fails to suggest, but teaches away from, 'at least one model for classifying the sound in the audio signal based on bandwidth,' etc." (Remarks page 16). However the examiner respectfully disagrees, and submits that the applicant's characterization of the goals in the system of ***Liu*** is too narrow. ***Liu*** discloses, "The phone level time resolution in our approach permits the algorithm to run quickly while maintaining the same accuracy as a frame level approach. Applying the new algorithms to a large sample of broadcast news programs resulted in improvements in speaker change detection accuracy, speech recognition accuracy, and speed" (Abstract). ***Liu*** is directed towards obtaining a fast speaker change detection system, without sacrificing accuracy. ***Liu*** discloses a system that attempts to reach a balance between processing speed and recognition and

speaker change detection accuracy. Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to modify the system in ***Liu*** in order to achieve that goal.

Applicant asserts that, "Liu directly teaches away from a non-phoneme based gender class" (Remarks page 20). However the examiner respectfully disagrees. The system of ***Liu*** uses a gender independent approach to the class decode, however ***Liu*** also discloses that, "gender difference would be easily detected with speaker change detection where not only the gender but also other speaker features are utilized to detect the difference. Doing so we can avoid using any complicated heuristic rules that may not be robust" (section 3 Phone-class decode). ***Liu*** specifically suggests a method to overcome the aforementioned drawbacks while still classifying based on gender difference.

Applicant's remaining arguments with respect to claims 1,10,16,23 and 28 have been considered but are moot in view of the new ground(s) of rejection.

Claim Rejections - 35 USC § 103

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject

matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

Claims 1-12, 14-22 and 28-31 are rejected under 35 U.S.C. 103(a) as being unpatentable over ***Liu*** ("Fast Speaker Change Detection for Broadcast News Transcription and Indexing" 1999) in view of ***Stanford*** (5,475,792).

As per claim 1, ***Liu*** discloses a method for classifying an audio signal containing speech information, the method comprising:

Receiving the audio signal (Figure 2);

Classifying a sound in the audio signal as a vowel class when a first phoneme-based model determines that the sound corresponds to a sound represented by a set of phonemes that define vowels (Section 3 Phone-Class Decode, paragraphs 3 and 4);

Classifying the sound in the audio signal as a fricative class when a second phoneme-based model determines that the sound corresponds to a sound represented by a set of phonemes that define consonants (Section 3 Phone-Class Decode, paragraphs 3 and 4; *fricatives are classified using a phoneme model. Since fricatives are by definition a specific type of consonant, it is inherent that they define consonants*); and

Classifying the sound in the audio signal based on at least one non-phoneme based model (Section 3 Phone-Class Decode, paragraph 3 and 4, *models are trained to classify non-speech, for example noise, music, laughter etc.*).

However, ***Liu*** does not disclose wherein the at least one non-phoneme based model including at least one model for classifying the sounds in the audio signal based on bandwidth. ***Stanford*** discloses a speech recognition system that enables recognition of high bandwidth or telephony (low bandwidth) speech signals (column 2 lines 30-32 and column 8 lines 36-44). ***Stanford*** discloses a speech recognition system that enables recognition of high bandwidth or telephony (low bandwidth) speech signals (column 2 lines 30-32 and column 8 lines 36-44). ***Stanford*** discloses that low bandwidth speech reduces the accuracy of speech recognizers (column 3 lines 37-39), and discloses a system that trains and uses a two separate codebook and phoneme models, one for low bandwidth speech and one for high bandwidth speech (column 8 lines 36-44) to increase the recognition accuracy of low bandwidth input. In addition, ***Liu*** attempts to provide a speaker change detection system which improves speaker change detection, speech recognition accuracy, and speed (Abstract).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to classify the sounds in the audio signal based on bandwidth in ***Liu***, since one of ordinary skill in the art has good reason to pursue the options within his or her technical grasp to achieve the predictable results of improving speech recognition accuracy while maintaining low processing time.

As per claim 2, ***Liu*** in view of ***Stanford*** disclose the method of claim 1, however ***Liu*** does not disclose a system wherein at least one non-phoneme based model includes models for classifying the sound in the audio signal based speaker gender. ***Liu*** does disclose that gender difference could be easily detected along with speaker change, where gender features as well as speaker features are used to tell the difference (Section 3 Phone-Class Decode, last paragraph).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to classifying the sound in the audio signal based on gender in ***Liu***, since one of ordinary skill in the art has good reason to pursue the options within his or her technical grasp to improve the accuracy of speaker change detection, while avoiding complicated heuristic rules that may not be robust, as indicated in ***Liu*** (section 3 Phone-Class decode, last paragraph).

As per claim 3 ***Liu*** in view of ***Stanford*** disclose the method of claim 1, and ***Liu*** further discloses wherein the at least one non-phoneme based model includes a model for classifying the sound in the audio signal as silence (Section 3 Phone-Class Decode, paragraph 5).

As per claim 4, ***Liu*** in view of ***Stanford*** disclose the method of claim 1, and ***Liu*** further discloses initially converting the audio signal into a frequency domain signal (section 4 Speaker Change Detection, *cepstral vectors are calculated and used for the distance measure. Cepstral vectors are derived from the spectrum of a signal; therefore the signal must have initially been converted into the frequency domain*).

As per claim 5, ***Liu*** in view of ***Stanford*** disclose the method of claim 1, and ***Liu*** further discloses generating cepstral features for the audio signal (section 4 Speaker Change Detection, *cepstral vectors are calculated and used for the distance measure).*

As per claim 6, ***Liu*** in view of ***Stanford*** disclose the method of claim 1, and ***Liu*** further discloses wherein the fricative class includes phonemes that relate to fricatives and obstruents (Section 3 Phone-Class Decode, paragraphs 3 and 4, *fricatives are a specific type of obstruent, therefore it is inherent that the class includes phonemes that relate to obstrunets*).

As per claim 7, ***Liu*** in view of ***Stanford*** disclose the method of claim 1, and ***Liu*** further discloses wherein the first and second phoneme-based models are Hidden Markov Models (Section 3 Phone-Class Decode, paragraph 6).

As per claim 8, ***Liu*** in view of ***Stanford*** disclose the method of claim 1, but ***Liu*** does not explicitly disclose classifying the sound in the audio signal as a coughing class when the sound corresponds to a non-speech sound. However, ***Liu*** does disclose coughing as a common non-speech sound (Section 2 Evaluation Metrics, second paragraph).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to classify the sound in the audio signal as a coughing class when the sound corresponds to a non-speech sound in ***Liu***, since classification of coughing as non-speech enables the exclusion of those frames during speaker clustering for identifying speakers, as taught by ***Liu*** (Section 3 Phone-Class Decode, first paragraph).

As per claim 9, ***Liu*** in view of ***Stanford*** disclose the method of claim 8, and ***Liu*** further discloses wherein the non-speech sound includes at least one of coughing, laughter, breath, and lip-smack (Section 2 Evaluation Metrics, second paragraph).

As per claim 10, ***Liu*** discloses a method of training audio classification models, the method comprising:

Receiving a training audio signal (Figure 2);

Receiving phoneme classes corresponding to the training audio signal (Section 3 Phone-Class Decode, paragraphs 3 and 4,*45 context-independent HMM phone models*

are trained, with models for vowels, fricatives, etc. It is inherent that phoneme classes corresponding to the training audio signal are received);

Training a first Hidden Markov Model (HMM), based on the training audio signal and the phoneme classes, to classify speech as belonging to a vowel class when the first HMM determines that the speech corresponds to a sound represented by a set of phonemes that define vowels (Section 3 Phone-Class Decode, paragraphs 3 and 4); and

Training a second HMM, based on the training audio signal and the phoneme classes, to classify speech as belonging to a fricative class when the second HMM determines that the speech corresponds to a sound represented by a set of phonemes that define consonants.

However, ***Liu*** does not disclose training at least one model to classify the sound based on a bandwidth of the sound. ***Stanford*** discloses a speech recognition system that enables recognition of high bandwidth or telephony (low bandwidth) speech signals (column 2 lines 30-32 and column 8 lines 36-44). ***Stanford*** discloses that low bandwidth speech reduces the accuracy of speech recognizers (column 3 lines 37-39), and discloses a system that trains and uses a two separate codebook and phoneme models, one for low bandwidth speech and one for high bandwidth speech (column 8 lines 36-44) to increase the recognition accuracy of low bandwidth input. In addition, ***Liu*** attempts to provide a speaker change detection system which improves speaker change detection, speech recognition accuracy, and speed (Abstract).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to classify the sounds in the audio signal based on bandwidth in ***Liu***, since one of ordinary skill in the art has good reason to pursue the options within his or her technical grasp to achieve the predictable results of improving speech recognition accuracy while maintaining low processing time.

As per claim 11, ***Liu*** in view of ***Stanford*** disclose the method of claim 10, and ***Liu*** further discloses wherein the phoneme classes include information that defines word boundaries (Section 5 Experiments and Results, Word-Error-Rate (WER), *the system determines the word error rate, or word recognition accuracy. Therefore it is inherent that the phoneme classes include information on word boundaries*).

As per claim 12, ***Liu*** in view of ***Stanford*** disclose the method of claim 11, and ***Liu*** further discloses wherein the method further comprises: receiving a sequence of transcribed words corresponding to the audio signal (Section 2 Evaluation Metrics, last paragraph, *reference transcription*); and generating the information that defines the word boundaries based on the transcribed words (Section 2 Evaluation Metrics, last paragraph, *the reference transcription is aligned with the acoustic data*).

As per claim 14, ***Liu*** in view of ***Stanford*** disclose the method of claim 10, however Liu does not disclose training at least one model to classify the sound based on gender of a speaker of the sound. ***Liu*** does disclose that gender difference could be easily detected

along with speaker change, where gender features as well as speaker features are used to tell the difference (Section 3 Phone-Class Decode, last paragraph).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to classifying the sound in the audio signal based on gender in ***Liu***, since one of ordinary skill in the art has good reason to pursue the options within his or her technical grasp to improve the accuracy of speaker change detection, while avoiding complicated heuristic rules that may not be robust, as indicated in ***Liu*** (section 3 Phone-Class decode, last paragraph).

As per claim 15, this claim has limitations similar to claim 6, and is therefore rejected for similar reasons.

As per claim 16, ***Liu*** discloses an audio classification device comprising:

A signal analysis component configured to receive an audio signal and process the audio signal by at least one of converting the audio signal to the frequency domain and generating cepstral features for the audio signal (section 4 Speaker Change Detection, *cepstral vectors are calculated and used for the distance measure. Cepstral vectors are derived from the spectrum of a signal; therefore the signal must have initially been converted into the frequency domain*).

A decoder configured to classify portions of the audio signal as belonging to at least one of a plurality of classes (Section 3 Phone-Class Decode, paragraphs 3 and 4), the classes including

A first phoneme-based class that applies to the audio signal when a portion of the audio signal corresponds to a sound, represented by a set of phonemes that define vowels (Section 3 Phone-Class Decode, paragraphs 3 and 4),

A second phoneme-based class that applies to the audio signal when a portion of the audio signal corresponds to a sound represented by a set of phonemes that define consonants (Section 3 Phone-Class Decode, paragraphs 3 and 4, *fricatives are classified using a phoneme model. Since fricatives are by definition a specific type of consonant, it is inherent that they define consonants*), and

At least one non-phoneme class (Section 3 Phone-Class Decode, paragraph 3 and 4, *models are trained to classify non-speech, for example noise, music, laughter etc.*).

However, ***Liu*** does not explicitly disclose whrein the decoder determines the at least one non-phoneme class using models that classify the portions of the audio signal based on bandwidth. ***Stanford*** discloses a speech recognition system that enables recognition of high bandwidth or telephony (low bandwidth) speech signals (column 2 lines 30-32 and column 8 lines 36-44). ***Stanford*** discloses that low bandwidth speech reduces the accuracy of speech recognizers (column 3 lines 37-39), and discloses a system that trains and uses a two separate codebook and phoneme models, one for low

bandwidth speech and one for high bandwidth speech (column 8 lines 36-44) to increase the recognition accuracy of low bandwidth input. In addition, ***Liu*** attempts to provide a speaker change detection system which improves speaker change detection, speech recognition accuracy, and speed (Abstract).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to classify the sounds in the audio signal based on bandwidth in ***Liu***, since one of ordinary skill in the art has good reason to pursue the options within his or her technical grasp to achieve the predictable results of improving speech recognition accuracy while maintaining low processing time.

As per claim 17, this claim contains limitations similar to those recited in claims 6 and 15, and is therefore rejected for similar reasons.

As per claim 18, this claim contains limitations similar to those recited in claim 7, and is therefore rejected for similar reasons.

As per claim 19, this claim contains limitations similar to those recited in claim 2, and is therefore rejected for similar reasons.

As per claim 20, this claim contains limitations similar to those recited in claim 3, and is therefore rejected for similar reasons.

As per claim 21, ***Liu*** in view of ***Stanford*** disclose the audio classification device of claim 16, and ***Liu*** further discloses wherein the plurality of classes additionally include: a third phoneme-based class that applies to the audio signal when a portion of the audio signal corresponds to a non-speech sound (Section 3 Phone-Class Decode, paragraph 3 and 4, *models are trained to classify non-speech, for example noise, music, laughter etc.*).

As per claim 22, this claim contains limitations similar to those recited in claim 9, and is therefore rejected for similar reasons.

As per claim 28, this claim contains limitations similar to those recited in claim 1, and is therefore rejected for similar reasons.

As per claim 29, this claim contains limitations similar to those recited in claim 4, and is therefore rejected for similar reasons.

As per claim 30, this claim contains limitations similar to those recited in claim 5, and is therefore rejected for similar reasons.

As per claim 31, this claim contains limitations similar to those recited in claim 8, and is therefore rejected for similar reasons.

Claims 23-27 are rejected under 35 U.S.C. 103(a) as being unpatentable over ***Liu*** in view of ***Stanford*** and further in view of ***Colbath*** ("Spoken Documents: Creating Searchable Archives from Continuous Audio" 2000).

As per claim 23, ***Liu*** discloses a system comprising:

Audio classification logic configured to classify the input audio data into at least one of a plurality of broad audio classes, the broad audio classes including a phoneme-based vowel class (Section 3 Phone-Class Decode, paragraphs 3 and 4), a phoneme-based fricative class (Section 3 Phone-Class Decode, paragraphs 3 and 4), and a non-phoneme based gender class (Section 3 Phone-Class Decode, paragraphs 3 and 4). ***Liu*** does not disclose an indexer configured to receive input audio data and generate a rich transcription from the audio data, the indexer including: a non-phoneme based bandwidth class, a speech recognition component configured to generate the rich transcription based on the broad audio classes determined by the audio classification logic, a memory system for storing the rich transcription, and a server configured to

receive requests for documents and respond to the requests by transmitting one or more of the rich transcriptions that match the requests.

Colbath discloses an indexer configured to receive input audio data and generate a rich transcription from the audio data (page 2, Component Technologies, first paragraph and page 4, System Architecture, first paragraph), a speech recognition component configured to generate the rich transcription based on the broad audio classes determined by the audio classification logic (page 2, Component Technologies, first paragraph), a memory system for storing the rich transcription, and a server configured to receive requests for documents and respond to the requests by transmitting one or more of the rich transcriptions that match the requests (page 4-5, System Architecture, server and browser).

In addition, ***Stanford*** discloses a speech recognition system that enables recognition of high bandwidth or telephony (low bandwidth) speech signals (column 2 lines 30-32 and column 8 lines 36-44). ***Stanford*** discloses that low bandwidth speech reduces the accuracy of speech recognizers (column 3 lines 37-39), and discloses a system that trains and uses a two separate codebook and phoneme models, one for low bandwidth speech and one for high bandwidth speech (column 8 lines 36-44) to increase the recognition accuracy of low bandwidth input. ***Liu*** attempts to provide a speaker change detection system which improves speaker change detection, speech recognition accuracy, and speed (Abstract).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to have an indexer configured to receive input audio data and generate

a rich transcription from the audio data, a speech recognition component configured to generate the rich transcription based on the broad audio classes determined by the audio classification logic, a memory system for storing the rich transcription, and a server configured to receive requests for documents and respond to the requests by transmitting one or more of the rich transcriptions that match the requests in ***Liu***, since it would create a system that integrates acoustic and linguistic technologies to construct a structural summary of continuous audio that is searchable by content, as indicated in ***Colbath*** (page 2, fourth paragraph).

It would also have been obvious to one of ordinary skill in the art at the time of the invention to classify the sound in the audio signal based on bandwidth in ***Liu***, since one of ordinary skill in the art has good reason to pursue the options within his or her technical grasp to achieve the predictable results of improving speech recognition accuracy while maintaining low processing time.

As per claim 24, ***Liu*** in view of ***Stanford*** further in view of ***Colbath*** disclose the system of claim 23, however ***Liu*** does not explicitly disclose wherein the broad audio classes further include a phoneme-based coughing class. ***Liu*** does disclose coughing as a common non-speech sound (Section 2 Evaluation Metrics, second paragraph).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to have the broad audio classes further include a coughing class in ***Liu***, since classification of coughing as non-speech enables the exclusion of those frames

during speaker clustering for identifying speakers, as taught by ***Liu*** (Section 3 Phone-Class Decode, first paragraph).

As per claim 25, ***Liu*** in view of ***Stanford*** further in view of ***Colbath*** disclose the system of claim 24, however ***Liu*** does not explicitly disclose wherein the coughing class includes sounds relating to coughing, laughter, breath, and lip-smack. ***Liu*** does disclose coughing, laughter, breath and lip-smack as a common non-speech sounds (Section 2 Evaluation Metrics, second paragraph).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to have the coughing class include sounds relating to coughing, laughter, breath, and lip-smack in ***Liu***, since classification of non-speech enables the exclusion of those frames during speaker clustering for identifying speakers, as taught by ***Liu*** (Section 3 Phone-Class Decode, first paragraph).

As per claim 26, ***Liu*** in view of ***Stanford*** further in view of ***Colbath*** disclose the system of claim 23, and ***Liu*** further discloses wherein the phoneme-based fricative class includes phonemes that define fricative or obstruent sounds (Section 3 Phone-Class Decode, paragraphs 3 and 4, *fricatives are a specific type of obstruent, therefore it is inherent that the class includes phonemes that relate to obstrunets*).

As per claim 27, ***Liu*** in view of ***Stanford*** further in view of ***Colbath*** disclose the system of claim 23, however neither ***Liu*** nor ***Stanford*** disclose wherein the indexer further includes at least one of: a speaker clustering component, a speaker identification component, a name spotting component, and a topic classification component. ***Colbath*** discloses wherein the indexer further includes at least one of: a speaker clustering component, a speaker identification component, a name spotting component, and a topic classification component (page 2, Component Technologies).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to have an indexer include at least one of: a speaker clustering component, a speaker identification component, a name spotting component, and a topic classification component in ***Liu*** and ***Stanford***, since it would create a system that integrates acoustic and linguistic technologies to construct a structural summary of continuous audio that is searchable by content, as indicated in ***Colbath*** (page 2, fourth paragraph).

Conclusion

Applicant's amendment necessitated the new ground(s) of rejection presented in this Office action. Accordingly, **THIS ACTION IS MADE FINAL**. See MPEP § 706.07(a). Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within

TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Dorothy Sarah Siedler whose telephone number is 571-270-1067. The examiner can normally be reached on Mon-Thur 9:30am-5:30pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Richemond Dorvil can be reached on 571-272-7602. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

DSS

TĀLIVALDIS IVARS ŠMITS
PRIMARY EXAMINER